

IN THE UNITED STATES DISTRICT COURT  
FOR THE DISTRICT OF OREGON

UNITED STATES OF AMERICA,

v.

RONALD CLAYTON RHODES,  
a/k/a “Big Fly,”  
LORENZO LARON JONES,  
a/k/a “Low Down,”

Defendants.

Case No. 3:19-cr-00333-MC

**OPINION AND ORDER**

**MCSHANE, District Judge.**

On August 4–5, 2022, this Court presided over an evidentiary hearing regarding Defendant Rhodes’s Motion for *Daubert* Hearing Regarding Admissibility of Toolmark Comparison Evidence, ECF 290, and Defendant Jones’s related Motion to Limit the Presentation of Ballistics Comparison Evidence by the Government, ECF 291. For the reasons set forth below, Defendants’ motions related to ballistic or toolmark comparison evidence, ECF 290 and ECF 291, are DENIED.

**STANDARDS**

Federal Rule of Evidence 702 governs the admissibility of expert testimony. It provides that:

A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:

- (a) the expert’s scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
- (b) the testimony is based on sufficient facts or data;
- (c) the testimony is the product of reliable principles and methods; and
- (d) the expert has reliably applied the principles and methods to the facts of the case.

“Under *Daubert*<sup>1</sup> and its progeny, including *Daubert II*<sup>2</sup>, a district court’s inquiry into admissibility is a flexible one.” *City of Pomona v. SQM N. Am. Corp.*, 750 F.3d 1036, 1043 (9th Cir. 2014) (citing *Alaska Rent-A-Car, Inc. v. Avis Budget Grp., Inc.*, 738 F.3d 960, 969 (9th Cir. 2013)). The trial court serves as “a gatekeeper, not a fact finder.” *Primiano v. Cook*, 598 F.3d 558, 565 (9th Cir. 2010) (internal quotation marks and citation omitted). The court “screen[s] the jury from unreliable nonsense opinions” but does not “exclude opinions merely because they are impeachable.” *Alaska Rent-A-Car*, 738 F.3d at 969–70 “The district court is not tasked with deciding whether the expert is right or wrong, just whether his testimony has substance such that it would be helpful to a jury.” *Id.*

Before admitting expert testimony into evidence, district court judges must determine whether the evidence is reliable and relevant under Rule 702. *Wendell v. GlaxoSmithKline LLC*, 858 F.3d 1227, 1232 (9th Cir. 2017). “Expert opinion testimony is relevant if the knowledge underlying it has a valid connection to the pertinent inquiry. And it is reliable if the knowledge underlying it has a reliable basis in the knowledge and experience of the relevant discipline.” *Primiano*, 598 F.3d at 565 (internal quotation marks and citation omitted). “Reliable expert testimony need only be relevant, and need not establish every element that the plaintiff must prove, in order to be admissible.” *Id.*

Determinations of the reliability of scientific expert testimony are guided by the factors outlined by the Supreme Court in *Daubert*. See *Daubert*, 509 U.S. at 593–95 (outlining the non-exclusive factors of (1) general acceptance in the scientific community, (2) peer review and

---

<sup>1</sup> *Daubert v. Merrell Dow Pharmas., Inc.*, 509 U.S. 579 (1993).

<sup>2</sup> *Daubert v. Merrell Dow Pharmas., Inc.*, 43 F.3d 1311 (9th Cir. 1995).

publication, (3) testability, and (4) error rate). Courts have recognized that this inquiry is flexible, and that these factors “neither necessarily nor exclusively appl[y] to all experts or in every case.” *Kumho Tire Co., Ltd. v. Carmichael*, 526 U.S. 137, 141 (1999). The court must consider whether the rate of error is sufficiently low, though the Ninth Circuit has said a methodology “need not be flawless in order to be admissible.” *United States v. Prime*, 431 F.3d 1147, 1153 (9th Cir. 2005). Finally, another “significant fact” is whether the expert is testifying “about matters growing naturally and directly out of research they have conducted independent of the litigation, or whether they have developed their opinions expressly for purposes of testifying.” *Daubert II*, 43 F.3d at 1317.

The *Daubert* standard of reliability applies not only to scientific testimony but also to testimony based on technical or other specialized knowledge. *Kumho Tire Co., Ltd.*, 526 U.S. at 141. But the Supreme Court also made clear that “the law grants a district court broad latitude when it decides *how* to determine reliability,” and that the *Daubert* factors “neither necessarily nor exclusively appl[y] to all experts or in every case.” *Id.* (emphasis in original); *see also id.* at 151 (explaining that certain factors may be more or less relevant). The Ninth Circuit has explained that when applying *Daubert* to non-scientific expert testimony, the district court “may consider the specific factors identified where they are reasonable measures of the reliability of proffered expert testimony,” but the court is not bound to “mechanically apply the *Daubert* factors.” *United States. v. Hankey*, 203 F.3d 1160, 1168 (9th Cir. 2000) (citation omitted).

## DISCUSSION

Defendants challenge the testimony of three Government experts with regard to firearm toolmark comparison evidence: Leland Samuelson and Shawn Malikowski, firearms and toolmark experts with the Oregon State Police (“OSP”) Forensic Laboratory; and Erich Smith, a

forensic examiner with the FBI Laboratory. ECF 272 at 1. Having reviewed the parties' briefing and considered the testimony and argument made at the *Daubert* hearing, this Court determines that the testimony of all three experts is admissible.

### **1. Testability**

Courts—including those that have excluded toolmark comparison evidence—have repeatedly found toolmark identification to be testable. *See United States v. Johnson*, Case No. (S5) 16 Cr. 281 (PGG), 2019 WL 1130258, at \*15 (S.D.N.Y. Mar. 11, 2019) (“There appears to be little dispute that toolmark identification is testable as a general matter.”); *United States v. Ashburn*, 88 F. Supp. 3d 239, 245 (E.D.N.Y. 2015) (“The AFTE methodology has been repeatedly tested.”); *United States v. Tibbs*, Case No. 2016-CF1-19431, 2019 WL 4359486, at \*7 (D.C. Super. Sep. 5, 2019) (“Although the NRC and PCAST reports have levied significant criticisms against firearms and toolmark analysis, courts have found that such reports do not affect the method’s testability. . . . [Toolmark analysis] can be, and ha[s] been, tested.”).

This Court acknowledges that at least one judge in this District has found that toolmark analysis was not testable for *Daubert* purposes. *See United States v. Adams*, 444 F. Supp. 3d 1248, 1260–64 (D. Or. 2020). In *Adams*, the court found that the proffered expert had failed to provide evidence that his toolmark analysis could be replicated. *Id.* at 1260 (“I do not, however, find that the AFTE comparison testing methodology, as described by [the proffered expert], is replicable.”). As a result, the court found in that case that the government did not satisfy the testability prong of *Daubert*. *Id.* at 1264. But this Court does not read *Adams* to say that the AFTE methodology or toolmark comparison is never testable as a matter of law; rather, the expert in *Adams* “could not define th[e] baseline in any objective way, nor could he explain the role it played in the actual comparison he made in this case.” *Id.* at 1261.

Here, on the other hand, Eric Smith, a Physical Scientist Forensic Examiner for the FBI, provided extensive testimony not only on AFTE methodology, but on the standards and training programs instituted at the FBI labs.<sup>3</sup> As part of that training, agents “look at thousands of comparisons of known matches and known non-matches.” Aug. 4, 2022 Tr. 14. Trainees appear before three oral boards, with “competency tests built into them,” testing the trainees’ knowledge and ability in comparing toolmark evidence. *Id.* at Tr. 15. After two years, the trainee may become a “qualified examiner.” Qualified examiners take yearly proficiency tests in firearms identification and toolmark identification. *Id.* at 17-18. These tests, performed by a third-party testing agency, evaluate both the examiner and the quality assurance system of the lab. The tests are treated exactly like ordinary evidence. *Id.* at 18. As the tests have “a known ground truth,” the examiner receives a full report on the accuracy of the results. *Id.* at 19.

In addition to these yearly proficiency tests, Smith testified to his personal involvement in numerous validation studies over the past two decades. Smith testified to approximately 25 validation studies over the past 25 years focusing on firearm toolmark examiner accuracy. *Id.* at 27-28. Many of these studies, most of which were “black box” studies, were specifically designed to test how often examiners *incorrectly* identified a toolmark as coming from a particular firearm; i.e., how often examiners returned a false positive. The Court agrees with the great weight of authority finding AFTE methodology can, and has been, tested. *See United States v. Chavez*, Case No. 15-CR-00285-LHK-1, 2021 WL 5882466, at \*2 (N.D. Cal. Dec. 13, 2021) (listing cases and noting the ““fact that numerous studies have been conducted testing the validity and accuracy of the AFTE method’ strongly suggests the method has and can be tested.””).

---

<sup>3</sup> Smith has a Bachelor’s of Science and a Master’s in Forensic Science from Virginia Commonwealth University.

Additionally, this Court does not measure replicability in the way proposed in *Adams*. In *Adams*, the court held that, “because [the AFTE methodology] cannot be explained in a way that would allow an uninitiated person to perform the same test in the same way that [the expert] did,” the methodology was not replicable or testable. *Adams*, 444 F. Supp. 3d at 1264. This focus on an uninitiated person is, in the Court’s view, misplaced. “Under *Daubert*’s testability factor, the primary requirement is that ‘[s]omeone else using the same data and methods . . . be able to replicate the result[s].’” *City of Pomona v. SQM N. Am. Corp.*, 750 F.3d 1036, 1047 (9th Cir. 2014) (citation omitted). Even if the expert in *Adams* failed to articulate the methodology in a way that satisfied the court in its gatekeeping function, this does not compel the finding that toolmark analysis be deemed untestable in all cases as a matter of law. Here, Smith testified that validation studies indicated examiners who completed a qualified training program tended to produce the same results. Aug. 4 Tr. 77-79. Training, as opposed to experience in the field, appears to be the critical factor. As Smith testified, “the training is critical.” *Id.* at 79. In other words, the studies indicated qualified examiners using the same data (i.e., toolmark evidence) and method (i.e., the AFTE Theory of Identification for sufficient agreement) generally replicate one another’s results.

The Court agrees with other courts that “The fact that there are subjective elements to the firearm and toolmark identification methodology is not enough to show that the theory is not ‘testable.’” *Chavez*, 2021 WL 58824666 at \*2 (citations omitted). When asked whether criticism of “sufficient agreement” is valid due to the methodology’s lack of objective datapoints or numeric thresholds, Smith testified:

No. . . . Because looking at the validation studies, we can definitely demonstrate the accuracy of examiners. So the fact that we can’t measure the toolmark and show a numerical value to support it does not take away from the accuracy of the examiner’s opinion. . . . The validation studies that have been reported give a high

degree of confidence on the performance of the examiners and their understanding of sufficient agreement.

*Id.* at 89. The Court agrees.

This Court finds that the AFTE methodology can be, and has been, tested. This factor weighs in favor of admissibility.

## 2. Error rate

Defendant primarily asserts that the error rate of toolmark analysis weighs against admission because it does not count inconclusive results as an error. This viewpoint is advanced primarily by defense experts Dean Faigman and Dr. Scurich.

Under the third *Daubert* factor, this Court considers “whether the technique has a known or potential rate of error.” *United States v. Romero-Lobato*, 379 F. Supp. 3d. 1111, 1119 (D. Nev. 2019). An error can either be a false negative—an “incorrect finding[] of dissimilarity”—or a false positive—an “incorrect affirmative identification.” *United States v. Mitchell*, 365 F.3d 215, 239 (3d Cir. 2004) (explaining error rates in fingerprint identification). In the context of firearm toolmark analysis, then, a false negative means that a certain bullet came from the tested firearm but the examiner concludes that it did not; conversely, a false positive means that a certain bullet *did not* come from the tested firearm, but the examiner concludes that it did. “[T]he focal point of the inquiry should be on the rate of false positives, as this is the type of error that could lead to a conviction premised on faulty evidence.” *Chavez*, 2021 WL 5882466, at \*4 (internal quotation marks omitted) (quoting *United States v. Harris*, 502 F. Supp. 3d 28, 39 (D.D.C. 2019)).

“[T]he weight of authority suggests the potential error rate [for toolmark analysis] is between 0–1%.” *Chavez*, 2021 WL 5882466, at \*4. Here, Smith testified extensively on numerous black box studies over the past two decades. Smith testified that “a fair conclusion” is

“that no matter the study design, no matter what the study conditions are, the false-positive rate is consistently around 1 percent or less.” Aug. 4 Tr. 146. It is true that courts that have found toolmark analysis is inadmissible (or must be limited) under *Daubert* have calculated the error rate slightly higher at 2.2 percent. *See Adams*, 444 F. Supp. 3d at 1264; *United States v. Shipp*, Case no. 19-cr-029-NGG, 422 F. Supp. 3d 762, 777–78 (E.D.N.Y. Nov. 26, 2019). But even this higher rate of error is still well short of the 5 percent error rate that the PCAST report identified as “an acceptable error rate from a scientific perspective. *Chavez*, 2021 WL 5882466, at \*4. While *Adams* focused rightly on the potential impact that false positive results may have on wrongful convictions, it does not address two additional layers of review: the second examiner in an accredited lab that reviews the first examiner’s results, and defense firearms experts who can provide a rebuttal to a government expert at trial. *Id.* “Thus, the underlying concern expressed in *Adams* and *Shipp*, the potential for a false conviction based on a false positive, appears to be much lower with the inclusion of another independent examiner, and even lower when there is a defense firearm expert. With just another independent examiner, the cumulative probability of a false positive rate could be as low as 0.05%.” *Id.* (footnote omitted). Smith testified that in the real world, due to quality assurance mechanisms in place with respect to FBI firearm and toolmark identifications, the chance of a false positive is lower than the 1% typically found in the various studies. Aug. 4, 2022 Tr. 178.

With respect to Dean Faigman and Dr. Scurich’s opinions that the error rates are unreliable because they do not include inconclusive results, this Court finds those opinions unsupported. Or, perhaps more accurately, the Court finds that while the relatively high rate of inconclusive results may be relevant to certain policy determinations—such as failing to eliminate a suspect from consideration—those concerns do not relate to the issue of paramount

importance to trial courts; i.e., the false positive rate which may result in a wrongful conviction. “[E]vidence of the false negative rate is often equivocal. While it might suggest a generally error-prone method, it is equally consistent with a very conservative method with a low false positive error rate. That is, a method may be designed to lower its false positive error rate by accepting a large number of false negatives out of an abundance of caution.” *United States v. Mitchell*, 365 F.3d 215, 239 n.19 (3rd Cir. 2004). Indeed, the *Tibbs* court, while excluding toolmark evidence, noted that Dr. Scurich’s opinion “that [inconclusive results] should be viewed as false positive errors (i.e., included among false identifications) . . . fails to make logical sense: while under laboratory conditions such inconclusives are surely some type of error, it does not follow that inconclusives are functionally the same as a false conclusion by an examiner who attributes a cartridge casing to a gun that did not fire it.” *Tibbs*, 2019 WL 4359486, at \*17. Put another way, while an inconclusive result is an error insofar as it means the methodology did not produce an answer, it is not an error in the sense that it falsely attributes a cartridge or casing to the wrong firearm.<sup>4</sup>

Even accepting the high error rates calculated by the *Adams* and *Shipp* courts, the error rate for toolmark analysis weighs strongly in the Government’s favor.

---

<sup>4</sup> Notably, Dean Faigman, formerly an advisor to PCAST, previously agreed with PCAST’s recommendation that toolmark studies should remove inconclusive results from the calculation. Aug. 4, 2022 Tr. 388. Although Dean Faigman testified he “made a mistake then,” this merely demonstrates that there exist legitimate reasons to not treat inconclusives as wrong answers. Similarly, the Government introduced numerous criticisms, from individuals even the defense experts agree are qualified to opine on the subject, of Dean Faigman and Dr. Scurich’s arguments regarding the treatment of inconclusives. *Id.* at 389-407. On this record, the vast majority of experts in the field—indeed, seemingly every expert other than Dean Faigman and Dr. Scurich—disagree with the defense experts as to the proper treatment of inconclusives in validation studies.

### 3. Peer review and publication

Notwithstanding Defendants' protestations, toolmark analysis has been subjected to peer review and publication. Still, because of the longstanding issues with the AFTE's peer review process and the relative recency with which AFTE has addressed concerns regarding its process, this Court finds that this factor does not weigh in favor of either party.

Defendants contend that “[u]ntil a couple of years ago, [the AFTE Journal's] peer review process suffered fundamental flaws by utilizing an open review process, failing to assure its published studies endured rigorous scientific scrutiny” and that “the AFTE does not freely allow the community of interested academics and scientists outside its own membership to access its publication,” which is behind a pay wall ECF 291 at 47–48. The Government responds that the AFTE has responded to criticisms of its process by “opening [the AFTE journal] up to outside members and now implementing a double-blind peer review process where applicable.” ECF 299 at 32–33 (citing *United States v. Cloud*, Case No. 1:19-cr-02032-SMJ, 2021 WL 7184484, at \*8 (E.D. Wash. Dec. 17 2021)). Like the *Cloud* court, this Court, “applauds the publication's changes and encourages AFTE . . . to continue to open their publications up for criticism and review from the larger scientific community.” *Cloud*, 2021 WL 7184484, at \*9.

While this Court is mindful that the AFTE Journal's peer review process may have been lacking and that the critical reports (e.g., the NCR Reports, PCAST Report, and validation studies) were published in peer review journals in their own right, most courts have not found that this factor weighs *against* admission. The *Shipp* court, which ultimately ruled that the toolmark analysis evidence was inadmissible, found that “even assigning limited weight to the substantial fraction of the literature that is published in the AFTE Journal, this factor still weighs in favor of admissibility.” 422 F. Supp. 3d at 776. Accordingly, this Court finds that this factor is neutral.

#### 4. Identifiable standards

*Daubert* directs courts to consider “the existence and maintenance of standards controlling [a] technique’s operation.” 509 U.S. at 594. Defendants challenge the standards to which toolmark analysts are held, arguing that the best-known non-match and sufficient agreement standards are too subjective to provide any guidance. ECF 291. This Court disagrees.

To begin, the “mere fact that an expert’s opinion is derived from subjective methodology does not render it unreliable.” *Romero-Lobato*, 379 F. Supp. 3d at 1120. As the Government points out, the subjectivity present in the best-known non-match and sufficient agreement standards is subject to industry standards that safeguard the process. ECF 299 at 34. These standards include: a specific laboratory’s standard operating procedures and guidelines; International Organization for Standardization (“ISO”)/International Electrotechnical Commission (“IEC”) Standard 17025; training, monitoring, validation of procedures, and regular proficiency testing to ISO/IEC Standard 17034. *Id.*

Moreover, forensic toolmark tests are subject to review by a second examiner who may either verify or disagree with the original result. *See Adams*, 444 F. Supp. 3d at 1266 (finding that this factor weighed in favor of admission and that toolmark comparison testing “was held to a high standard and subject to quality control”). Additionally, Smith testified to the FBI’s requirement that the examiner make detailed notes and photographs documenting the examiner’s process. Aug. 4, 2022 Tr. 106.

Our reports are very long. A report has the results, the opinions that I produced; but along with it, you are going to have the methods and the limitations that go along with each of the examinations. That’s purposefully done so, if I’m not present to testify and my report is submitted into evidence, the jury will have an opportunity to understand, if they read my report, what was done.

*Id.* at 106-07.

This factor favors admissibility.

## 5. Acceptance

Finally, Defendants urge this Court to find that the AFTE methodology does not enjoy widespread acceptance in the scientific community. ECF 291 at 50–51. Specifically, Defendants point to the “committees of scientific experts assembled for the NRC and PCAST reports,” whom Defendants argue “uniformly and unequivocally conclude that firearms/toolmarks examination does not enjoy general acceptance.” *Id.* at 52.

But the weight of authority suggests that the AFTE method does enjoy general acceptance in the relevant scientific community—forensic ballistic examiners. *See Romero-Lobato*, 379 F. Supp. 3d at 1122 (“The AFTE method certainly satisfies [the general acceptance] element.”); *Ashburn*, 88 F. Supp. 3d. at 247 (“The AFTE theory . . . has been widely accepted in the forensic community.”); *United States v. Otero*, 849 F. Supp. 2d. 425, 435 (D. N.J. 2012) (noting that even courts critical of the AFTE method have found that it is “widely accepted among examiners as reliable”); *United States v. Monteiro*, 407 F. Supp. 2d 351, 372 (D. Mass. 2006). Smith testified that firearm and toolmark identification is accepted within both the toolmark and broader forensic science communities. Aug. 4, 2022 Tr. 111. The Court agrees. The few courts that have defined the scientific community more broadly miss the mark. *See Tibbs*, 2019 WL 4359486, at \*21 (citing the NRC and PCAST reports as evidence that the “wider academic and scientific community does not generally accept this theory”). But these reports only concluded that more studies were needed, and indeed more studies have been done. *See* ECF 299-15, Ex. O (2018 AFTE Journal study); ECF 299-16, Ex. P (2020 Journal of Forensic Sciences study); ECF 299-17, Ex. Q (2021 FBI study).

This Court acknowledges Defendants’ arguments that the community of forensic examiners have a vested interest in the continued acceptance of forensics. But these arguments go to the weight of the evidence; they are not indicia that the relevant community does not

generally accept the AFTE method. Although not dispositive, this favor weighs in favor of admissibility.

#### **6. Conclusion**

An analysis of the *Daubert* factors, in light of the parties' briefings and exhibits, as well as the testimony and arguments made at the hearings on August 4–5, 2022, compels a finding that the toolmark comparison evidence is admissible.<sup>5</sup> Neither the Government nor its experts have shied away from the fact that toolmark comparison has a subjective element. Still, the method is testable, has a low error rate, is based on identifiable standards, and is accepted in the relevant scientific community. This Court finds that the AFTE Journal's peer review process has undergone substantial improvements in recent years, but also notes that the journal had previously been inaccessible to non-members; this factor cannot be said to weigh in favor of any party more than the other.

This is not to say that Defendants have not articulated some warranted skepticism with respect to toolmark analysis. But this skepticism is proper fodder not for the outright exclusion of evidence on *Daubert* grounds, but rather for robust cross-examination at trial.

#### **CONCLUSION**

Defendants' motions related to ballistic or toolmark comparison evidence, ECF 290 and ECF 291, are DENIED.

**IT IS SO ORDERED.**

DATED this 17th day of January, 2023.

/s/ Michael McShane  
Michael J. McShane  
United States District Judge

---

<sup>5</sup> The Government stipulated that Smith would testify only to his belief that certain bullets came from a certain gun. FBI regulations prohibit Smith from testifying that the identification was to the exclusion of all other firearms.